UPDATE AND REVISION IN A DYNAMIC EPISTEMIC FRAMEWORK

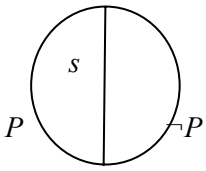
Johan van Benthem, Amsterdam & Stanford, http://staff.science.uva.nl/~johan/

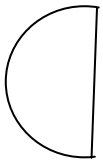
Games and Verification, Newton Institute, Cambridge, 3 July 2006

Abstract *Real games are not just bare processes: they involve information update as moves are played, belief revision as expectations change, and perhaps even preference change as goals keep shifting*. Logical Dynamics is about making such rich (inter-)actions first-class citizens in logical systems. Dynamic-epistemic logics treat information update from observed events, changing the current doxastic -epistemic model. Belief revision can be treated in the same format, using update rules for plausibility relations which have been proposed also for preference change. We obtain complete sets of axioms for particular revision mechanisms, and standard modal frame correspondence analysis for revision postulates in a general AGM style. `Ref.` http://staff.science.uva.nl/~johan/DL-BR-new.pdf.

1 Information update as systematic model change

Card examples: learning P eliminates the worlds where P is false. In a picture:

from $\boldsymbol{M}$ 

to $\boldsymbol{M}|P\,s$ 

Epistemic content of questions/answers involves iterated and common knowledge.

2 Dynamic epistemic logic: compositional analysis of effects

Static epistemic logic Language $p \mid \neg\phi \mid \phi\vee\psi \mid K_i\phi \mid C_G\phi$, models $\boldsymbol{M} = (W, \{\sim_i \mid i\in G\}, V)$, with worlds W, accessibility relations $\sim_i$, and valuation V. Crucial epistemic truth conditions: $\boldsymbol{M}, s \models K_i\phi$ iff for all t with $s \sim_i t$: $\boldsymbol{M}, t \models \phi$, and $\boldsymbol{M}, s \models C_G\phi$ iff for all t that are reachable from s by some finite sequence of $\sim_i$ steps ($i\in G$): $\boldsymbol{M}, t \models \phi$.

Dynamic logic of public announcement *PAL*: add action expressions: $!P$ for all formulas P, and modal operators describing their effects (one simultaneous recursion):

$\boldsymbol{M}, s \models [!P]\,\phi$ iff *if* $\boldsymbol{M}, s \models P$, *then* $\boldsymbol{M}|P, s \models \phi$

Theorem 1 (Plaza, Gerbrandy) PAL without C_G is axiomatized completely by

the usual laws of epistemic logic plus the following *reduction axioms*:

$[!P]q \leftrightarrow P \rightarrow q$ *for atomic facts* q

$[!P]\neg\phi \leftrightarrow P \rightarrow \neg[!P]\phi$

$[!P]\phi\wedge\psi \leftrightarrow [!P]\phi \wedge [!P]\psi$

$[!P]K_i\phi \leftrightarrow P \rightarrow K_i(P \rightarrow [!P]\phi)$

$[!P][!Q]\phi \leftrightarrow [!(P \wedge [!P]Q)]\phi$

Methodology Add dynamic superstructure to static base logic. *Compositional* analysis of all post-conditions. Requires design for *pre-encoding* in static language. Example: $[!P]C_G\phi$

requires new notion: 'conditional common knowledge' with reduction axiom:

$$[!P]C_G(\phi, \psi) \leftrightarrow C_G(P \wedge [!P]\phi,\ [!P]\psi).$$

By-product: reduction dynamic logic to completeness/validity in static language. ***Program*** 'Dynamification' of existing logics, making the underlying actions explicit.

3 Update by general events with partial observation

Email: epistemic-dynamic function of *cc*, *bcc*. Computer security. Current high-light: *Games* designed to manipulate information flow (*Cluedo*). Partial observation of events.

Event models A = $(E, \{\sim_i \mid i \in G\}, \{PRE_e \mid e \in E\})$. Scenario: relevant events, relations $\sim_i$ encode what agents cannot distinguish. I check my card: you cannot tell 'my seeing red' from 'my seeing black'. Events *e* have *preconditions* PRE_e for their execution: my having a red card, not knowing the answer to my question, etc. Update Rule: for any epistemic model $(\boldsymbol{M}, s)$ and event model $(\boldsymbol{A}, e)$, the ***product model*** $(\boldsymbol{M}\mathbf{x}\boldsymbol{A}, (s, e))$ has

Domain $\{(s, e) \mid s$ a world in $\boldsymbol{M}$, e an event in $\boldsymbol{A}$, $(\boldsymbol{M}, s) \models PRE_e\}$,
Accessibility: $(s, e) \sim_i (t, f)$ iff *both* $s \sim_i t$ *and* $e \sim_i f$,
Valuation for atoms *p* at (s, e) is that at *s* in $\boldsymbol{M}$. (can be generalized to world change)

Product update deals with misleading actions as well as truthful ones, and with *belief* as well as knowledge. Epistemic models can even get *larger* as update proceeds (*bcc*)!

Dynamic-epistemic logic LEA: $p \mid \neg\phi \mid \phi\vee\psi \mid K_i\phi \mid C_G\phi \mid [\boldsymbol{A}, a]\phi$: $(\boldsymbol{A}, e)$ for any event model with actual event *e*. Semantics, key clause: $\boldsymbol{M}, s \models [\boldsymbol{A}, e]\phi$ iff $\boldsymbol{M} \mathbf{x} \boldsymbol{A}, (s, e) \models \phi$.

Theorem 2 (BMS) *LEA* is effectively axiomatizable and decidable.

The key reduction axiom is the one extending that for public announcement:

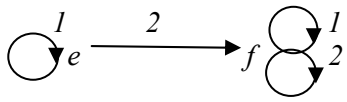
$$[\boldsymbol{A}, e]K_i\phi \leftrightarrow PRE_e \rightarrow \wedge\{K_i[\boldsymbol{A}, f]\phi)) \mid f \sim_i e \text{ in } \boldsymbol{A}\}$$

Further issues: extensions to richer languages like ***LCC***, or *epistemic μ–calculus.*

Needed for dealing with common knowledge.

(a) ***Idealized agents***. Product Update: *Perfect Recall*, and *No Miracles*. Diversity?

(b) ***Common knowledge*** or ***common belief*** in subgroups: *secrets*. Example *(BEK):*



4 Belief revision theory, the bare necessities

Update $T+A$, revision $T*A$, contraction $T–A$. *AGM Postulates*: general constraints. Grove sphere models. Representation theorems. Key topics: $B \in T*A$ close to (but not quite) Ramsey test for $T \vdash A \Rightarrow B$, issue of iteration, proliferation of revision policies. *Conditional logic*. Restrictions: only factual assertions *A*, not epistemic ones – single agent scenarios – recent view: special case of 'belief merge' from different sources.

5 Belief change under hard facts

Knowledge and belief Reinterpret $K_i\phi$ to weaker $[]_i\phi$: 'to the best of i's information'. Even better: have both knowledge and belief modalities in base language. Example a model in whose actual world x, I believe that $\neg p$, though p is in fact the case:

$$p, x \xrightarrow{\cdots\cdots} y, \neg p$$

Problem with update *DEL* does not do true belief revision. 'Hard announcement' event $!p$ turns this into the one-world model $\{x\}$ where I believe that p, but even $B\bot$!

Belief and plausibility models Solution: conditional logic of relative plausibility:

$\boldsymbol{M}, s \models B_i\phi$ iff $\boldsymbol{M}, t \models \phi$ for all worlds t *minimal* in the ordering $\lambda xy.\ \leq_{i,s} xy$.

Belief change under hard facts:

$$[!P]\ B_i\phi \quad \leftrightarrow \quad P \rightarrow B_i([!P]\phi \mid P)$$

Conditional belief helps *pre-encode* beliefs we would have if we learnt certain things:

$\boldsymbol{M}, s \models B_i(\phi/\psi)$ iff $\boldsymbol{M}, t \models \phi$ for all t minimal for $\lambda xy.\ \leq_{i,s} xy$ in $\{u \mid \boldsymbol{M}, u \models \psi\}$.

Satisfies the standard principles of the minimal conditional logic.

Theorem 3 The logic of conditional belief under public announcements is axiomatized completely by (a) any complete base logic of $B_i(\phi/\psi)$ for favorite model class, (b) *PAL* reduction axioms, plus (c) a reduction axiom for conditional beliefs:

$$[!P]\ B_i(\phi/\psi) \quad \leftrightarrow \quad P \rightarrow B_i([!P]\phi \mid P \wedge [!P]\psi)$$

Discussion: close to Ramsey test, but difference in 'dynamics': explain role of $[!P]$. Combines immediately with earlier knowledge axioms. *(Preserve interplay axioms?)*

6 Belief change under soft facts

Soft triggers Call for belief revision $*p$ 'softer' than elimination, introducing just a 'preference' for p-worlds. Spohn, Veltman: *default rule* $A \Rightarrow B$ does not say all A-worlds are B. Just makes 'exceptional' $A \wedge \neg B$-worlds less plausible. 'Soft information' does not eliminate, it *changes the plausibility order* of worlds. Recurrent instruction:

Lexicographic upgrade $\Uparrow P$ changes the current model $\boldsymbol{M}$ to $\boldsymbol{M}\Uparrow P$:
P-worlds now better than all $\neg P$-worlds; within zones, old order remains.

Social revolution: underclass P now becomes upper class. Other policies (Rott's '27'; or Macchiavelli: 'conservative belief revision', $\uparrow P$, just co-opt *leaders* of the underclass!) Dynamic language in standard *DEL*-style:

$\boldsymbol{M}, s \models [\Uparrow P]\phi$ iff $\boldsymbol{M}\Uparrow P, s \models \phi$.

The static pre-encoding is again done by *conditional beliefs*:

Theorem 4 The dynamic logic of lexicographic upgrade is axiomatized completely by logic of conditional belief + compositional analysis of effects of revision:

$[\Uparrow P]\, q \leftrightarrow q, \qquad [\Uparrow P]\, \neg\phi \leftrightarrow \neg[\Uparrow P]\phi, \qquad [\Uparrow P]\, (\phi\wedge\psi) \leftrightarrow [\Uparrow P]\phi \wedge [\Uparrow P]\psi$

$[\Uparrow P]\, B(\phi|\psi) \leftrightarrow (E(P \wedge [\Uparrow P]\psi) \wedge B([\Uparrow P]\phi \mid P \wedge [\Uparrow P]\psi)) \vee B([\Uparrow P]\phi \mid [\Uparrow P]\psi)$

Here E is the existential modality 'in some world'. Case of just factual assertions:

$[\Uparrow P]\, B(\phi|\psi) \leftrightarrow (E(P \wedge \psi) \wedge B(\phi \mid P \wedge \psi)) \vee B(\phi \mid \psi)$

$[\Uparrow P]\, B\phi \leftrightarrow (E\, P \wedge B([\Uparrow P]\phi \mid P)) \vee B[\Uparrow P]\phi$ (a bit more 'Ramsey-like)

Conclusion *Constructive belief revision theory can be studied by standard modal techniques.* We find complete axiomatizations for concrete belief revision policies defined as systematic relation change (e.g., conservative upgrade $\Uparrow P$). Many further policies axiomatized at *ILLC* in the past year, also including 'point assignment', etc.

7 *AGM* postulates as modal frame correspondence

Abstract constraints on model-changing operations are just modal frame correspon-dences (as with '*K4* – transitivity'). Let $\Box A$ be any operation from models $\boldsymbol{M}$ and sets of worlds A in it to a new model $\boldsymbol{M}\Box A$ with the same worlds but a changed relation $\leq_s$.

Fact The formula $[\Box p]Bp$ says that the best worlds in $\boldsymbol{M}\Box p$ are all in p.

Fact The formula $B(q|p) \leftrightarrow [\Box p]Bq$ expresses 'rule by the upper classes'.

Now *invert* earlier completeness results to see the content of reduction axioms:

Theorem 5 Eliminative update fixed by the earlier $[\Box p]\, K\, q \leftrightarrow (p \rightarrow K\, [\Box p]\, q)$.

Theorem 6 The formula $[\Box p]\, B(q|r) \leftrightarrow (E(p \wedge r) \wedge B(q \mid p \wedge r) \vee B(q \mid r))$

holds in a frame iff the operation interpreting $\Box p$ is lexicographic upgrade.

Similar analysis for most exciting *AGM* Postulates (*two* model changing operations):

(a) $[\Box(p\wedge q)]B\, r \rightarrow [!q][\Box p]B\, r$

(b) $[\Box p]E\, q \wedge [!q][\Box p]B\, r \rightarrow [\Box(p\wedge q)]B\, r$

Conclusion *Axiomatic belief revision theory can be studied by standard modal means.*

But of course, we would really want extended *AGM* postulates for *conditional* belief! Connected to problem of iterated belief revision: what about $[\Uparrow A]\, [\Uparrow B]\, \phi$?

8 Discussion

Richer triggers *DEL* event models model much richer multi-agent scenarios, where agents have beliefs about events. Graded *strengths* of beliefs (Spohn, Aucher, Liu).

Real world change easily added (van Benthem, van Eijck & Kooi). Katsuno-Mendelzon.

Obstacle to a happy marriage *DEL* 'backward-looking': computes what we believe after an event takes place, via *preconditions*. *AGM*: 'forward-looking' instructions "come to believe",

"see to it that". General events via postconditions ill-defined. Possible remedy: work over temporal universes which constrain future developments.

9 Longer-term processes: conversation, games, learning theory

Program structures in conversation: Sequential composition ***;***, Guarded choice *IF THEN ELSE*, Guarded iteration *WHILE DO*. Even parallel composition **||** makes sense.

Games Strategic interaction (learning/teaching games: 'Teaching the Unwilling').
On top of standard game theory, *DEL* gives fine-structure of deliberation and moves.

Epistemic temporal logics Branching time: Halpern et al., Parikh et al. For connections across the area, see van Benthem & Pacuit 2006, *'The Tree of Knowledge in Action'*.

Process algebra *DEL* imports bit of model construction, ' epistemic process algebra'.

Learning theory Natural continuation of update and revision logic 'by other means'.

Related approaches Segerberg, Girard, van Ditmarsch, Baltag & Smets, *ESSLLI 2005*.

References Stanford 2006 seminar page http://staff.science.uva.nl/~johan/seminar2006.html has papers on all of the above + comparisons with other frameworks. See also the papers 'Research' at http://staff.science.uva.nl/~johan/ with a lot of further references to the literature.